7Л

УДК 004.82

В.В. Любченко
Одесский национальный политехнический университет, г. Одесса, Украина
vira.lyubchenko@gmail.com

Формализация оценки достаточного уровня декомпозиции модели предметной области

В статье сформулирован принцип достаточности декомпозиции. Предложен количественный показатель для измерения степени рациональности декомпозиции, выполнено модельное исследование его поведения.

Введение

При построении моделей предметной области методом декомпозиции исследуемая предметная область условно разделяется на ряд подобластей (компонент, элементов). Следует понимать, что детальность декомпозиции предметной области на компоненты зависит от уровня знаний о предметной области и целевого назначения ее модели.

Для структуризации процесса построения модели целесообразно выполнять декомпозицию итеративно, что дает возможность анализировать иерархическую структуру и представить ее в виде дерева декомпозиции. В этом случае компонентам предметной области $C=\{c_1,\ldots,c_n\}$ соответствует множество вершин дерева $V=\{v_1,\ldots,v_n\}$ (каждая вершина $v_i \in V$ представляет отдельный компонент $c_i \in C$), связанных между собой множеством ребер $E=\{e_1,\ldots,e_m\}$.

В дереве декомпозиции ребра могут соединять только вершины, принадлежащие разным уровням. При этом ребро между вершинами v_i и v_j соответствует отношению есть-часть (is-a-part) между соответствующими компонентами, то есть

$$\exists e\left(v_i, v_j\right) \& \, depth\left(v_i\right) > depth\left(v_j\right) \rightarrow c_i \subset c_j , \qquad (1)$$

где $depth(v_i)$ – длина пути от вершины v_i к корневой вершине.

Объединив вершины, находящиеся на одной глубине, можно сформировать уровневые множества вершин дерева декомпозиции. Каждое уровневое множество соответствует определенному уровню детализации представления в модели предметной области.

Так как процедура декомпозиции выполняется итеративно, то важным вопросом является выбор критерия остановки. Очевидно, что с увеличением глубины декомпозиции растет информативность получаемой структуры. Но при этом возрастает и сложность получаемой структуры. Можно предположить, что на определенном этапе продолжение выполнения декомпозиции станет нерациональным, так как информация о предметной области приобретет большую подробность, чем это необходимо. Это приведет к получению значительной сложности дерева декомпозиции, увеличению трудоемкости выполнения декомпозиции или росту значений других характеристик сложности. То есть должен существовать некоторый оптимальный с точки зрения соотношения показателей информативность-сложность уровень детализации, при достижении которого становится нецелесообразным продолжение итеративного выполнения декомпозиции.

Целью данной работы является определение критерия, позволяющего выбрать оптимальную глубину дерева декомпозиции.

Принцип достаточности декомпозиции

Анализ ситуации показывает, что для всех компонент предметной области существует достаточный уровень детализации, который обеспечивает их конструктивное использование. Степень достаточности определяется исходя из целей построения модели и ожидаемых результатов ее использования. Чаще всего этот процесс протекает на интуитивном уровне, и эксперт не фиксирует в достаточной мере внимание на принципах, которыми руководствуется. Тем не менее, этот момент является одним из важнейших при формировании модели предметной области. Таким образом, появляется необходимость в формализации данного принципа – принципа достаточности декомпозиции.

Сформулируем *принцип достаточности декомпозиции*: всегда существует достаточный (не избыточный) уровень детализации модели предметной области, который может быть задан на основе формализованного анализа целей и задач использования этой модели.

Принцип достаточности декомпозиции можно рассматривать как развитие принципа существенности [1] и принципа достаточности [2], которые используют при формировании моделей сложных объектов.

Рассмотрим, каким образом можно оценить рациональность декомпозиции модели предметной области.

Показатель рациональности декомпозиции

В [3] показано, что для измерения информативности неориентированных графов можно использовать показатель информационной емкости, основанный на значениях степеней вершин

$$I = \sum_{i=1}^{N_v} a_i \log_2 a_i \,, \tag{2}$$

где $N_v = |V|$ – количество вершин графа, a_i – степень i-й вершины, т.е. количество инцидентных ей ребер.

В [3] показано, что мерой сложности неориентированного графа может служить значение максимальной энтропии для этого графа, которое вычисляется по формуле

$$H_{\max} = A \log_2 A \,, \tag{3}$$

где $A = \sum_{i=1}^{N_v} a_i$ – смежность графа.

Тогда в качестве показателя рациональности декомпозиции можно использовать отношение

$$K_R = {I}/{H_{\max}} \,, \tag{4}$$

где I, $H_{\max}$ определяются формулами (2) и (3) соответственно. Справедливо следующее утверждение.

Утверждение. При увеличении глубины дерева декомпозиции значение коэффициента рациональности декомпозиции уменьшается.

Для доказательства справедливости утверждения необходимо показать, что значение показателя информационной емкости I с увеличением глубины дерева декомпозиции растет медленнее значения максимальной энтропии H_{max}, т.е. $\Delta I<\Delta H_{max}$ или $\Delta H_{max} - \Delta I > 0$.

$$\Delta H_{max} - \Delta I = \left(H_{max}^2 - H_{max}^1\right) - \left(I^2 - I^1\right) = \left(H_{max}^2 - I^2\right) - \left(H_{max}^1 - I^1\right). \quad (5)$$

Шеннон определил информацию как меру сокращения энтропии для системы с постоянным количеством элементов [4]:

$$I = H_{max} - H, \quad (6)$$

где H – остаточная энтропия системы. Используя (6), можем записать (5) как

$$\left(H_{max}^2 - I^2\right) - \left(H_{max}^1 - I^1\right) = H^2 - H^1 = \Delta H. \quad (7)$$

Известно, что увеличение количества элементов всегда должно увеличивать полную энтропию системы, т.е. $\Delta H>0$, поэтому действительно $\Delta H_{max} - \Delta I > 0$.

Рассмотрим поведение показателя рациональности при увеличении глубины дерева декомпозиции. Модельные исследования проводились для бинарного и тернарного дерева с изменением значения глубины от 3 до 10. Результаты исследований приведены в табл. 1.

Таблица 1 – Зависимость показателя рациональности декомпозиции от глубины дерева декомпозиции

Глубина дерева декомпозиции L	Значение показателя рациональности K_r	
	для бинарного дерева	для тернарного дерева
3	0,268	0,261
4	0,227	0,206
5	0,193	0,167
6	0,168	0,140
7	0,148	0,122
8	0,132	0,105
9	0,119	0,093
10	0,108	0,084

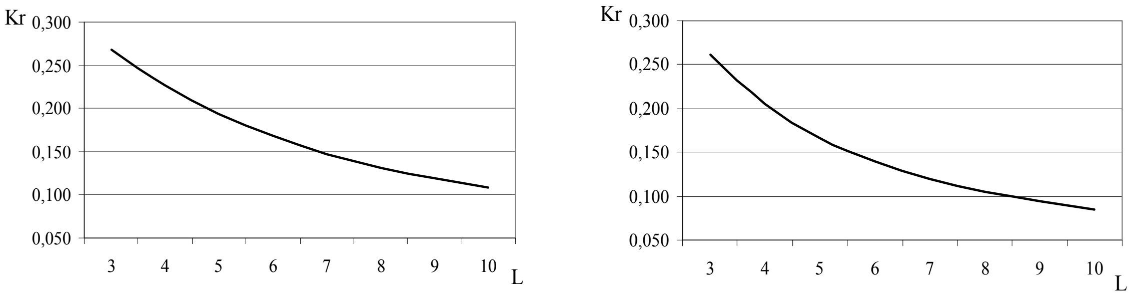
Как видно из таблицы, модельный эксперимент подтверждает утверждение о том, что при увеличении глубины дерева декомпозиции значение показателя рациональности декомпозиции уменьшается. Аппроксимирующие зависимости для показателя рациональности декомпозиции K_r в случае регулярных деревьев имеют схожий характер:

$$K_r = \begin{cases} \dfrac{1,27}{1,65+L}, & \textit{если дерево бинарное} \\ \dfrac{1,20}{1,64+L} - 0,02, & \textit{если дерево тернарное} \end{cases}, \quad (8)$$

где L – глубина дерева декомпозиции.

На рис. 1 показано, как изменяется показатель рациональности декомпозиции с увеличением глубины дерева.

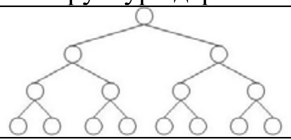
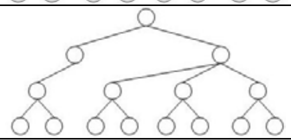
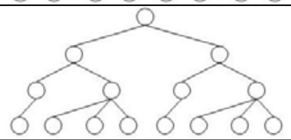
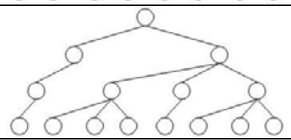
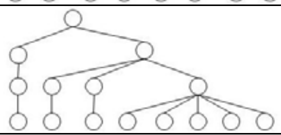
Необходимо обратить внимание на то, что значение показателя рациональности декомпозиции зависит не только от количества уровней декомпозиции и количества вершин на каждом из уровней, но также зависит и от структуры соединений между вершинами, принадлежащими разным уровням. Пример влияния нарушений структуры дерева приведен в табл. 2.



а) б)

Рисунок 1 – Графики изменения значений K_r для бинарного (а) и для тернарного (б) дерева. L – глубина дерева декомпозиции

Таблица 2 – Численные показатели для разных структур дерева декомпозиции глубины 4

Структура дерева	I	H_{max}	K_r
	30,529	134,606	0,227
	31,020	134,606	0,230
	31,510	134,606	0,234
	32,000	134,606	0,238
	33,510	134,606	0,249

Как и следовало ожидать, чем больше нарушается регулярность структуры соединений, тем больше становится значение показателя рациональности декомпозиции.

Нарушение структуры деревьев приводит к тому, что аппроксимирующие зависимости для показателя рациональности декомпозиции K_r не имеют схожий характер. Например, в случае искажения структуры бинарного дерева, были получены такие зависимости:

$$K_r = \begin{cases} \dfrac{1,27}{1,65+L}, & \textit{если дерево бинарное без искажений} \\ 0,07+\exp(-0,23L-0,89), & \textit{если дерево с регулярными искажениями} \\ 0,21+\dfrac{4,13}{36,20+\exp(L)} & \textit{если дерево с нерегулярными искажениями} \end{cases}. \quad (9)$$

Однако тенденция уменьшения показателя рациональности декомпозиции K_r с увеличением ее глубины сохраняется (рис. 2).

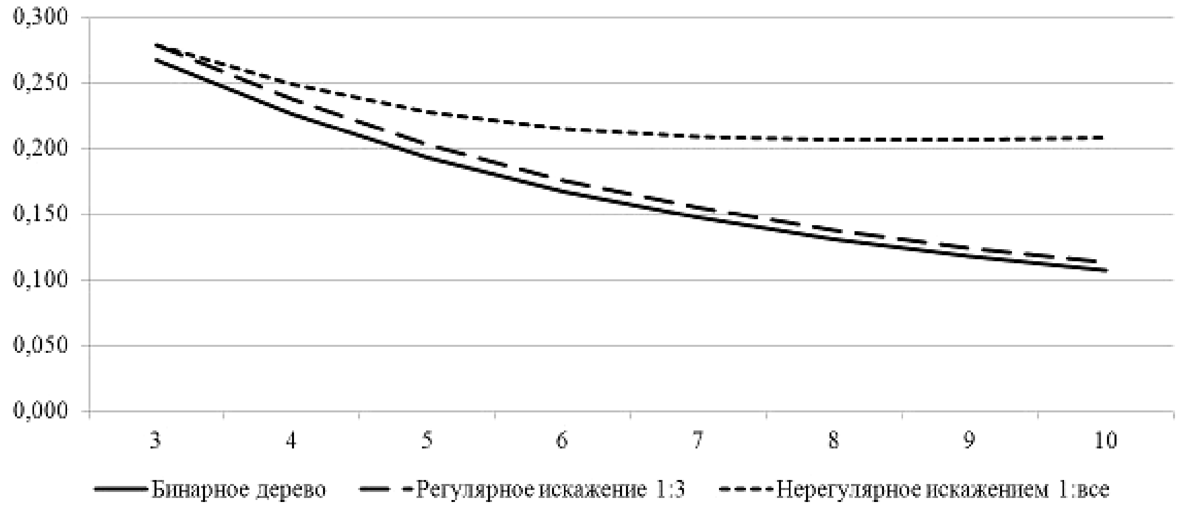


Рисунок 2 – Графики изменения значений K_r
в зависимости от глубины дерева декомпозиции L

Критерий достаточности декомпозиции

Свойство уменьшения значения показателя рациональности с увеличением глубины дерева декомпозиции позволяет сформулировать критерий остановки для итеративной процедуры декомпозиции модели предметной области.

Можно утверждать, что всегда найдется такое значение $K_{\min}$, для которого увеличение глубины декомпозиции (то есть $K_r<K_{\min}$) приведет к необоснованной детализации. Это означает, что детальность описания предметной области растет, но с большой вероятностью модель будет в дальнейшем использоваться не в полном объеме.

С другой стороны, всегда существует такое значение $K_{\max}$, для которого уменьшение глубины декомпозиции (то есть $K_r>K_{\max}$) не обеспечивает детализации, необходимой для достижения целей и задач использования модели предметной области.

На базе общего принципа достаточности декомпозиции становится возможным сформулировать три основных условия, которые являются формализованными критериями целенаправленного формирования модели предметной области:

1) условие недостаточной детализации модели:

$$K_r > K_{\max}; \qquad (10)$$

2) условие избыточной детализации модели:

$$K_r < K_{\min}; \qquad (11)$$

3) условие конструктивной декомпозиции:

$$K_{\min} \le K_r \le K_{\max}. \qquad (12)$$

Выводы

Подводя итог, следует отметить, что принцип достаточности декомпозиции обосновывает необходимость введения количественного показателя рациональности декомпозиции.

Свойство уменьшения значения показателя рациональности декомпозиции с увеличением степени детализации декомпозиции позволяет сформулировать критерий

остановки для итеративной процедуры декомпозиции. Таким критерием может быть достижение заданного априори критического значения показателя. Использование данного критерия позволяет на этапе формализации представления предметной области корректно задать модель предметной области в виде дерева ее декомпозиции.

Литература

1. Математическое моделирование химических производств / [Кроу К., Гамилец А., Хофман Т. и др.]. – М. : Мир, 1973. – 392 с.
2. Крисилов В.А. Информационная технология принятия решений в задачах АСУ на базе количественной интегральной оценки сложных объектов : дис. ... доктора техн. наук / Крисилов В.А. – Одесса : ОНПУ, 2004. – 242 с.
3. Bonchev D. Shannon's Information and Complexity / D. Bonchev // Mathematical Chemistry Series. – 2003. – Vol. 7, Complexity in Chemistry. – P. 155-187.
4. Шеннон К.Э. Работы по теории информации и кибернетике / Шеннон К.Э. – М. : Иностр. лит., 1963. – 829 с.

В.В. Любченко

Формалізація оцінки достатнього рівня декомпозиції моделі предметної області

У статті сформульовано принцип достатності декомпозиції. Запропоновано кількісний показник для вимірювання ступеня раціональності декомпозиції, проведено модельне дослідження його поведінки.

V. Liubchenko

Formalization of Sufficient Level Evaluation for Domain Model Decomposition

The principle of decomposition sufficiency is formulated. The quantitative measure for rationality degree of decomposition is offered, the model research of its behavior is executed.

Статья поступила в редакцию 28.05.2010.